

Teoría de las comunicaciones

Primer cuatrimestre de 2017

Departamento de Computación
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Trabajo práctico 1: Wiretapping

Integrante	LU	Correo electrónico
??	???	??@???.???
??	???	??@???.???
??	???	??@???.???

Reservado para la cátedra

Instancia	Docente	Nota
Primera entrega		
Segunda entrega		

Comentarios del corrector:

Índice

1. Introducción	3
2. Métodos y ambientes	4
3. Red hogareña	4
4. Red laboral	6
5. Conclusiones	7

Índice de figuras

1. Información red hogareña	5
2. Visualización red hogareña	5
3. Información red laboral	6
4. Visualización red laboral	7

1. Introducción

En este informe nos proponemos analizar diferentes redes locales utilizando herramientas derivadas de la teoría de la información. El objetivo es usar datos triviales de los paquetes que pasan por la red para descubrir patrones y topologías sin conocer previamente nada sobre los nodos de la red. Para lograr esto utilizamos direcciones de la capa de enlace en un principio y luego datos de paquetes ARP (address resolution protocol, un protocolo utilizado para obtener una dirección de enlace dada una dirección de red para que 2 nodos en una misma red local de acceso compartido puedan comunicarse) para obtener mediciones de información y entropía de la misma red modelada de modo diferente.

2. Métodos y ambientes

Las redes serán analizadas con dos herramientas que escuchan un medio compartido en modo promiscuo. La primera herramienta modela la red como una fuente S cuyos posibles símbolos son $S_{unicast}$ si el paquete leído es enviado a un nodo específico y $S_{broadcast}$ si el paquete es enviado a la dirección de broadcast ($ff : ff : ff : ff : ff : ff$), ningún paquete es filtrado, cualquier paquete que llega nuestro enlace será tenido en cuenta.

La segunda herramienta modelará la red como una fuente $S1$ cuyos posibles símbolos son todas las direcciones de red disponibles en la red (como esto es desconocido, vamos a asumir que todos los símbolos observados son todos los símbolos posibles) en paquetes ARP que nos encontramos en la red. La consigna era definir una función que designe algún símbolo como distinguido fundado en algún resultado matemático y para esto distinguimos a aquellos símbolo o símbolos que tengan información menor a la entropía dado que estos serán más comunes de ver (por la definición de información y entropía), es decir, que los otros nodos de la red piden su dirección de enlace muy frecuentemente. Esto hará distinguir a nodos internos muy usados o bien al gateway por defecto de la red por la cual todos los nodos salen a otras redes como internet.

Las N redes que elegimos para experimentar son las siguientes

- Red hogareña: red local wi-fi en una casa particular, lo único que sabemos es que contiene una cantidad muy limitada de computadoras personales y dispositivos móviles.
- ??
- ??

3. Red hogareña

Esta red en principio es muy controlada y solo sabemos que cuenta con un número pequeño de nodos. La red usa en su mayoría wi-fi por lo cual es de esperar que el nodo que usamos para monitorear la red reciba amplia cantidad de paquetes who-has y is-at dado que no hay switches segmentando la red que filtren respuestas is-at de terceros.

El primer experimento, el cual ve la red como una fuente con 2 símbolos únicamente, arroja luego de monitorear la red durante 30 minutos que en ella la mayor parte del tráfico es unicast y solo el 17% es broadcast.

	frecuencia	información
broadcast	0.17	2.56
unicast	0.83	0.27

Estos valores nos dan una entropía de 0.66 bits (siendo el máximo 1 dado que hay 2 símbolos en principio equiprobables) lo cual concuerda con las expectativas dado que en una red local hogareña se espera ver más que nada tráfico unicast entre los nodos y el gateway hacia internet.

En el segundo experimento, el cual modela la red basado en la dirección a resolver en mensajes ARP, usamos la misma captura utilizada en el experimento anterior. El programa que usamos para analizar los resultados destacó 3 nodos de la red cuando esperábamos solo 1. El siguiente gráfico muestra la información de cada símbolo y la entropía de la fuente, como se puede ver hay 3 nodos destacados bajo nuestra definición.

Una investigación más meticulosa concluyó que 192.168.1.1 es el gateway por defecto de la red mientras que los otros 2 nodos destacados, 192.168.1.11 y 192.168.1.101 ni siquiera existen en la red, las direcciones no fueron asignadas y ningún nodo las está usando más un nodo de la red está constantemente intentando resolver su dirección de enlace. Solo podemos suponer que este host tiene

Información para cada símbolo de la fuente

--- Entropía

Símbolo	Información
192.168.1.101	2.0
192.168.1.11	2.1
192.168.1.1	2.4
192.168.1.117	3.2
192.168.1.116	3.5
169.254.255.255	4.6
192.168.1.114	5.4
192.168.1.111	5.4
192.168.1.113	5.8
12.240.200.9	5.8
192.168.1.104	7.2

Las interacciones de los nodos se puede ver en más detalle con el siguiente grafico donde se puede ver a los nodos destacados y las relaciones con los otros nodos. Notar que en el grafico, así como en el experimento también aparece un nodo con dirección 196.254.255.255 la cual es la dirección que el sistema operativo windows asignado a un nodo cuando este no puede contactar a ningún servidor DHCP para que le asigne una dirección libre. Usualmente los nodos operan con esta dirección durante segundos o minutos hasta que se puedan proveer de una dirección valida.

5

4. Red laboral

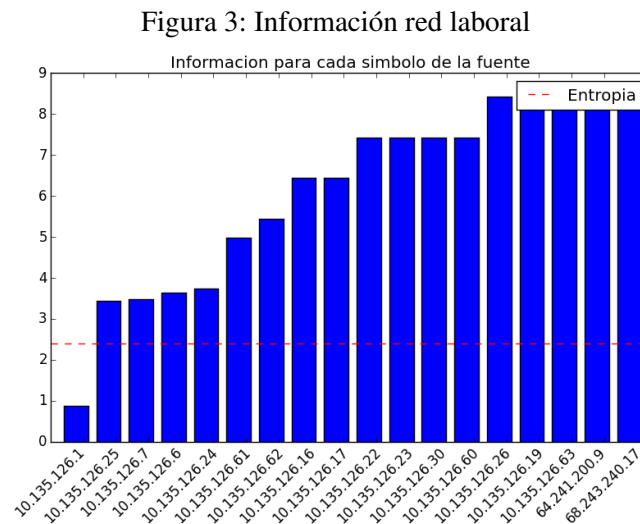
Para explorar nuevas topologías y relaciones decidimos monitorear una red laboral en una oficina de mediana envergadura. En esta red local conviven nodos de empleados, servidores de distinto uso, así como impresoras de red y otros dispositivos de oficina.

El primer experimento, el cual ve la red como una fuente con 2 símbolos únicamente, monitoreo la red local a través de un enlace ethernet durante 20 minutos. Como era de esperarse cerca del 100% del tráfico es unicast dado que es una red con mucho tráfico hacia un servidor en particular o hacia el gateway.

	frecuencia	información
broadcast	0.03	5.06
unicast	0.97	0.04

Estos valores nos dan una entropía de 0.19 bits (siendo el máximo 1 dado que hay 2 símbolos en principio equiprobables), dado que el tráfico broadcast sin ser despreciable solo es un porcentaje pequeño del tráfico total.

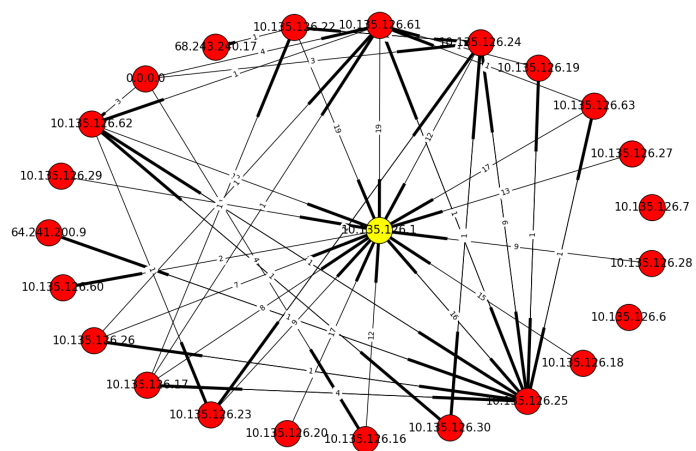
En el segundo experimento, el cual modela la red basado en la dirección a resolver en mensajes ARP, usamos la misma captura utilizada en el experimento anterior. Bajo nuestra definición para destacar nodos (que la información del símbolo de ese nodo sea menor o igual a la entropía de la fuente) solo un nodo fue destacado por amplio margen, el gateway por defecto de la red.



La visualización de la red modelada como la fuente descriptiva muestra una topología y relación entre nodos muy similar a la realidad en términos de tráfico total y disposición.

Un análisis manual del tráfico mostró 2 fenómenos de ARP no usuales, ARP gratuitos y ARP de sondeo. Los ARP gratuitos pueden ser tanto who-has como is-at donde la dirección de origen como destino son las mismas (y la dirección de enlace destino es broadcast). El uso de los mismos es precargar o refrescar las tablas de otros nodos para evitar tener que traducir en tiempo real una dirección de red. Los ARP de sondeo tienen un fin similar, un nodo al que se le asigna manualmente o automáticamente una dirección de red, envía un ARP con dicha dirección a la red y si algún nodo le responde entonces sabrá que la dirección ya está en uso y evitara la colisión de alguna forma. Ningún de estos tipos de mensajes están oficialmente documentados pero forman parte de toda red que necesite autoregularse correctamente.

Figura 4: Visualización red laboral



5. Conclusiones

TODO